Kinetic Schemes for Saint-Venant Equations with Source Terms on Unstructured Grids

Emmanuel Audusse, Marie-Odile Bristeau, Benoît Perthame

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Kinetic schemes for Saint-Venant equations with source terms on unstructured grids

Emmanuel Audusse , Marie-Odile Bristeau , Benoit Perthame

No 3989

Juillet 2000

THÈME 4

ISSN 0249-6399 ISRN INRIA/RR--3989--FR+ENG

INRIA
ROCQUENCOURT

Kinetic schemes for Saint-Venant equations with source terms on unstructured grids

Emmanuel Audusse *, Marie-Odile Bristeau †, Benoit Perthame ‡

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet M3N

Rapport de recherche n ° 3989 — Juillet 2000 — 44 pages

Abstract: We consider the Saint-Venant (or Shallow Water) system which is an usual model to describe the flows in rivers or coastal areas. This hyperbolic system of conservation laws is solved on unstructured meshes by a kinetic scheme based on a finite volume approach. An important property of this scheme is the preservation of the water height positivity when applications with dry areas are considered. Following some hypothesis an entropy inequality is proved.

The standard kinetic scheme is modified to deal with varying bed slope and particularly to preserve equilibrium states such as still water. Moreover the source terms due to the arbitrary bottom topography have to be discretized in such a way to balance the flux gradients for these equilibriums.

We illustrate the properties of the scheme on different test cases for which exact solutions are available and on more realistic applications.

Key-words: Saint-Venant system, shallow water system, unstructured grid, finite volume method, kinetic schemes, source terms, equilibrium scheme.

(Résumé : tsvp)

* M3N
† M3N, Email : Marie-Odile.Bristeau@inria.fr
‡ M3N, Email : Benoit.Perthame@ens.fr

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
Téléphone : 01 39 63 55 11 - International : +33 1 39 63 55 11
Télécopie : (33) 01 39 63 53 30 - International : +33 1 39 63 53 30

Schémas cinétiques pour les équations de Saint-Venant avec termes sources sur des maillages non structurés

Résumé : On considère le système des équations de Saint-Venant qui modélise les écoulements en "eaux peu profondes". Ce système hyperbolique de lois de conservation est résolu sur des maillages non structurés par un schéma cinétique qui s'appuie sur une approche volumes finis. Une propriété importante de ce schéma est d'assurer la positivité de la hauteur d'eau dans les cas où le domaine de calcul inclus des zones sèches. Sous certaines hypothèses, on peut montrer une inégalité d'entropie.

Le schéma cinétique standard est modifié pour traiter les fonds variables et, en particulier, pour que le schéma préserve des états d'équilibre tels que l'"eau au repos". De plus, on définit la discrétisation des termes sources dus à la topographie du fond pour préserver cet équilibre.

On illustre les propriétés du schéma sur différents problèmes tests pour lesquels on connaît la solution analytique et sur des applications plus réalistes.

Mots-clé : Système de Saint-Venant, eaux peu profondes, maillages non structurés, volumes finis, schémas cinétiques, termes sources, schémas d'équilibre.

1 Introduction

We consider the Saint-Venant system, a particular case of shallow-water equations, which is an usual model to describe, for instance, the flows in rivers or coastal areas. This system can be derived from the three dimensional incompressible Euler system with a free surface, using several assumptions (small height of water, hydrostatic pressure, vertical homogeneity of horizontal velocities ...). It allows to describe the flow at time $t \geq 0$, and at the point $x \in \mathbb{R}^2$, through the height of water $h(t,x) \geq 0$, and its velocity $u(t,x) \in \mathbb{R}^2$, by the hyperbolic system

$$\frac{\partial h}{\partial t} + \mathrm{div}(hu) = 0, \tag{1.1}$$

$$\frac{\partial hu}{\partial t} + \mathrm{div}(hu \otimes u) + \nabla(\frac{g}{2}h^2) + gh\nabla Z = 0, \tag{1.2}$$

where g denotes the gravity intensity and $Z(x)$ the bottom height, and therefore $h+Z$ is the level of water surface (see Fig.(1.1)). We denote also $q(t,x) = h(t,x)u(t,x)$, the discharge.

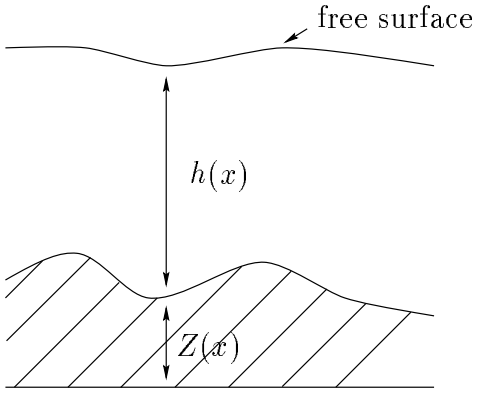

Figure 1.1:

Usually, several other terms are added to this system in order to take into account frictions on the bottom and the surface and other physical features but here we neglect all these phenomena.

The unidimensional version of Saint-Venant equations is well adapted for ideal rectangular rivers and we use it to introduce our methods before extending it into two dimensions situations. It is written

$$\frac{\partial h}{\partial t} + \frac{\partial hu}{\partial x} = 0, \tag{1.3}$$

$$\frac{\partial hu}{\partial t} + \frac{\partial}{\partial x}(hu^2 + \frac{gh^2}{2}) + gh\frac{\partial Z}{\partial x} = 0. \tag{1.4}$$

Notice however that this system corresponds to a simple case but a more general equation can be stated for rivers with variable sections.

The difficulty to understand the convergence of numerical schemes is related to the deep mathematical structure of such hyperbolic system; the first existence proof of weak solutions after shocks in the large is due to [18] in 1995. It is based on the kinetic interpretation of the system which is also a method to derive numerical schemes with good properties.

A classical approach for solving hyperbolic systems consists in using finite volume schemes, see Godlewski and Raviart [10], LeVeque [16]. This method can be used on general triangular grids with a finite element data structure using a particular control volume which is the mediane based dual cell as it was introduced several years ago for aerodynamics problems (see [2] for instance). Let us recall that finite volume methods require to compute fluxes at the control volumes interfaces, and the overall stability of the method (and more precisely *entropy property*) requires some upwinding in the interpolation of the fluxes, this is usually called the 'building block' or 'approximate Riemann solver', see [1] for an application to shallow water equations.

But none of these approximate Riemann solvers is able to provide all the desirable properties of a 'good' scheme, and the main difficulties arising specifically with the Saint-Venant system is that the scheme preserves equilibrium states like the steady state given by still water ($h + Z = Cst$, $u = 0$). This difficulty was pointed out by several authors, see Greenberg and Leroux [14], LeVeque [17]. Approaches to build stable schemes which preserve equilibrium have been proposed recently. The Roe solver [25], has been modified in order to preserve equilibriums in Bermudez and Vasquez [4], Bermudez et al [3] and the difficulty to allow h to vanish is treated in [20]. Also, a recent approach by LeVêque [17] is based on using the Godunov scheme for an appropriately extended system, and an approach based on centred schemes is used in [26]. For the simpler case of a scalar conservation law, a kinetic scheme that preserves equilibrium and which is proved to be convergent, is presented in Botchorishvili *et*

al [5], and another approach based on non-conservative products has been developed by Gosse and Leroux [12], Gosse [11]. Finally, we refer to Godinaud et al [9] for a new approach where the notion of Riemann solver is replaced by a more general concept.

In this paper we describe a possible resolution of the system (1.1), (1.2) using a *kinetic solver* and present a variant which has the property to preserve the still water equilibrium. Also it preserves the very simplicity of the usual kinetic solver and its stability properties. Especially, it is well known to be able to treat vacuum ($h = 0$ here, it corresponds to dry soils, then the system losses hyperbolicity) and satisfies precise in-cell entropy inequalities. We refer to Perthame [22] for a survey of the theoretical properties of these schemes, Perthame and Qiu [23] for the actual implementation in two dimensions and Bouchut [6] for a general construction for 2x2 systems. Their main drawback is the lack of accuracy on steady contact discontinuities (they are computed exactly by Godunov, Osher-Engquist or Roe solvers) which makes them unable to compute boundary layers or shear flows in classical fluid dynamics problems. For Saint-Venant equations this problem does not exist however, since the only possible discontinuities are shocks, and therefore the kinetic schemes might be a better compromise between accuracy, stability and efficiency.

Another advantage of kinetic schemes is that it allows to derive a scheme for the advection of temperature $T(t,x)$ (or a concentration of pollutant) which is both monotone in T and conservative in hT. This additional, and decoupled equation is written in a conservative form as

$$\frac{\partial hT}{\partial t} + \mathrm{div}(huT) = 0, \tag{1.5}$$

but *monotinicity* is better seen on the developed form

$$\frac{\partial T}{\partial t} + u \cdot \nabla T = 0. \tag{1.6}$$

We do not consider that point in this paper but an analysis can be found in [7].

After presenting the kinetic formulation of the shallow water system in Section 2, we explain the 1D standard kinetic scheme and we establish its properties in Section 3, we present the associated numerical results in Section 4. Then, we introduce a variant which is adapted to the still water equilibrium in Section 5 and we treat its extension for 2D problems in Section 6. The Section 7 illustrates the possibilities of the method by numerical results.

2 Properties of Saint-Venant system and kinetic formulation

We recall here some well known properties of the shallow water system, and by analogy with Euler equations for compressible gas, we link the macroscopic Saint-Venant system to a microscopic description of the fluid which is the departing point of the methods developed in this paper.

2.1 Entropy inequality

We recall two main properties of the shallow water system. First of all, the system admits an invariant region $h(t,x) \geq 0$, and the water height h can indeed vanish (flooding zones, dry regions, tidal flats). This is related to a theoretical and numerical difficulty because the system losses hyperbolicity at $h = 0$. Another fundamental property is the entropy property. We call *entropy solution* to the Saint-Venant system, a weak solution which satisfies the entropy inequality given in the following theorem.

Theorem 2.1 *The systems (1.1)-(1.2) and (1.3)-(1.4) are strictly hyperbolic for $h > 0$. They admit a mathematical entropy (which is also the energy)*

$$E(h,u,Z) = h\frac{|u|^2}{2} + \frac{gh^2}{2} + ghZ, \tag{2.1}$$

which satisfies

$$\frac{\partial}{\partial t}(E) + \mathrm{div}[u\,(E + \frac{gh^2}{2})] \leq 0. \tag{2.2}$$

We do not prove this theorem which relies on the classical theory of hyperbolic equations and simple algebraic calculation, see Serre [27], Dafermos [8]. We just recall that for smooth solutions the inequality in (2.2) is an equality. In particular the system admits a family of steady states given by the relations

$$h\,u = \mathrm{C}_1, \tag{2.3}$$

$$\frac{|u|^2}{2} + gh + gZ = \mathrm{C}_2, \tag{2.4}$$

where C_1 and C_2 are two arbitrary constants.

2.2 Kinetic formulation

We now explain the kinetic formulation of this system. For reason of simplicity we introduce it for the one-dimensional problem but the construction is very similar in two dimensions (see Sec. 6). Let us introduce a real function χ defined on $\mathbb{R}$ which has the following properties

$$\begin{aligned} &\chi(-w) = \chi(w) \geq 0, \\ &\int_{\mathbb{R}} \chi(w)dw = 1, \qquad \int_{\mathbb{R}} w^2\chi(w)dw = 1, \end{aligned} \tag{2.5}$$

and let us construct the density of particles $M(t,x,\xi)$ defined by a so-called *Gibbs equilibrium*

$$M(t,x,\xi) := M(h,\xi-u) = \frac{h(t,x)}{\tilde{c}}\ \chi(\frac{\xi-u(t,x)}{\tilde{c}}), \tag{2.6}$$

with

$$\tilde{c}^2 = \frac{g\,h}{2}. \tag{2.7}$$

With these definitions we write a kinetic representation of the Saint-Venant system and we have the following theorem:

Theorem 2.2 *The functions (h,u) are weak solutions to the Saint-Venant system (1.3)–(1.4) if and only if $M(h,\xi-u)$ satisfies the kinetic equation*

$$\frac{\partial M}{\partial t} + \xi\frac{\partial M}{\partial x} - g\frac{\partial Z}{\partial x}\frac{\partial M}{\partial \xi} = Q(t,x,\xi), \tag{2.8}$$

for some "collision term" $Q(t,x,\xi)$ which satisfies for a.e. t, x,

$$\int_{\mathbb{R}} Q\ d\xi = 0, \qquad \int_{\mathbb{R}} \xi Q\ d\xi = 0. \tag{2.9}$$

It is an entropy solution if additionally

$$\int_{\mathbb{R}} \xi^2 Q\ d\xi \leq 0. \tag{2.10}$$

Proof of Theorem 2.2. The proof relies on a very obvious computation. The two Saint-Venant equations, and the entropy inequality (2.2) are equivalent with the equation (2.8) once integrated in ξ against 1, ξ and ξ^2. The lefthand sides coincide exactly thanks to the relations (2.5) and the righthand sides vanish or are nonpositive according to (2.9) and (2.10). These are consequences of the easy relations deduced from (2.5):

$$\begin{pmatrix} h \\ hu \\ hu^2 + \frac{g\ h^2}{2} \end{pmatrix} = \int_{\mathbb{R}} \begin{pmatrix} 1 \\ \xi \\ \xi^2 \end{pmatrix} M(h, \xi - u)\ d\xi, \tag{2.11}$$

$$E(h, u, Z) = \int_{\mathbb{R}} (\frac{\xi^2}{2} + gZ)\ M(h, \xi - u)\ d\xi. \tag{2.12}$$

□

This theorem has a very important consequence: the non-linear shallow water system can ve viewed as a single linear equation on a nonlinear quantity M, for which it is easier to find a simple numerical scheme with good theoretical properties.

Notice however that this Theorem is a very weak statement. Especially in the *kinetic formulations* of isentropic gas dynamics (which are more general than the mere Saint-Venant system), all the family of weak entropy inequalities are satisfied, not just one, and the righthand side Q vanishes for smooth solutions, not only two moments, see [19]. This formalism has the drawback to lead to more complex schemes.

To conclude this subsection, we present a first example of χ functions. The simplest possible choice is

$$\chi(\omega) = \begin{cases} \frac{1}{2\sqrt{3}} & \text{for} |\omega| \leq \sqrt{3} \\ 0 & \text{otherwise.} \end{cases} \tag{2.13}$$

Another natural choice is presented later.

3 The standard kinetic scheme for 1D Saint-Venant equations

We now pass to the description of *kinetic schemes* for the Saint-Venant system. In a first subsection we recall some notations for the finite volume method on a general

mesh (see [10]). In a second subsection we use the above kinetic formulation to introduce a kinetic scheme and we show some fundamental properties like positivity and in-cell entropy inequality in the next subsections.

3.1 The 1D finite volume method

We use a mesh of $\mathbb{R}$ which vertices are denoted X_i, $i \in \mathbb{Z}$. The control volume (cell) is defined as $C_i = [X_{i-\frac{1}{2}}, X_{i+\frac{1}{2}}]$ with $X_{i+\frac{1}{2}} = (X_{i+1} + X_i)/2$. For simplicity we suppose that we have a regular mesh and we denote Δx the length of a cell C_i. We also use a timestep Δt, and set $t^n = n\ \Delta t$.

To present the general finite volume method, we consider a system of first order conservation laws without source terms:

$$\frac{\partial U}{\partial t} + \frac{\partial F(U)}{\partial x} = 0. \tag{3.1}$$

For Saint-Venant system (1.3)–(1.4) with $Z = Cst$, we have:

$$U = \begin{pmatrix} h \\ \\ q \end{pmatrix}, \quad F(U) = \begin{pmatrix} q \\ \\ \dfrac{q^2}{h} + \dfrac{gh^2}{2} \end{pmatrix}.$$

We denote by U_i^n the cell average of the exact solution at time t^n

$$U_i^n = \frac{1}{\Delta x} \int_{C_i} U(t^n, x) dx. \tag{3.2}$$

We integrate the equation (3.1) on the set $C_i \times (t^n, t^{n+1})$, and, integrating by parts the divergence term, we obtain

$$U_i^{n+1} - U_i^n + \frac{\Delta t}{\Delta x}[F_{i+\frac{1}{2}}^n - F_{i-\frac{1}{2}}^n] = 0 \tag{3.3}$$

with

$$F_{i+\frac{1}{2}}^n = \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} F(U(t, X_{i+\frac{1}{2}})) dt. \tag{3.4}$$

To obtain a finite volume scheme it remains to find a numerical value $F_{i+\frac{1}{2}}^n$ as an interpolation of the component of the flux $F(U)$ on the interface at the point $X_{i+\frac{1}{2}}$. We will use the kinetic formulation to define an approximation of this flux (we use the same notations for the approximated quantities).

3.2 Standard Kinetic scheme

The *kinetic scheme* is an upwind interpolation in order to determine the numerical fluxes $F^n_{i+\frac{1}{2}}$ in (3.3), based on the kinetic formulation introduced in Section 2. We assume that the bottom topography Z is given at each node.

We assume that we know the solution U^n_i at time t^n for each node, we define the following algorithm with three steps:

- We define $M^n_i = M(h^n_i, \xi - u^n_i)$ with $M(h, \xi - u)$defined by (2.6).

- We use the microscopic equation (2.8). Since this equation is linear we can apply a simple upwind scheme with a method of characteristics for the force term. It leads to the following discretisation which defines a density function $f^{n+1}_i(\xi)$:

$$f^{n+1}_i(\xi - g\Delta t Z'_i) - M^n_i(\xi) + \frac{\Delta t}{\Delta x}\xi\ [M^n_{i+\frac{1}{2}}(\xi) - M^n_{i-\frac{1}{2}}(\xi)] = 0, \tag{3.5}$$

where $Z'_i = \frac{Z_{i+1} - Z_{i-1}}{2\Delta x}$ and the flux $M^n_{i+\frac{1}{2}}(\xi)$ is computed by an upwind formula

$$M^n_{i+\frac{1}{2}}(\xi) = \begin{cases} M^n_i(\xi) & \text{for } \xi \geq 0, \\ M^n_{i+1}(\xi) & \text{for } \xi \leq 0. \end{cases}$$

The density function $f(\xi)$ is not an equilibrium (see remark (3.1)).

- We define U^{n+1}_i by

$$U^{n+1}_i = \int_{\mathbb{R}} \begin{pmatrix} 1 \\ \xi \end{pmatrix} f^{n+1}_i(\xi)\ d\xi. \tag{3.6}$$

□

Then, noticing that

$$\int_{\mathbb{R}} \xi\ f^{n+1}_i(\xi - g\Delta t Z'_i)\ d\xi = q^{n+1}_i + g\Delta t\ h^{n+1}_i Z'_i,$$

we obtain from (3.5),(3.6) the macroscopic scheme

$$U^{n+1}_i - U^n_i + \sigma[F^n_{i+\frac{1}{2}} - F^n_{i-\frac{1}{2}}] + \Delta t\ S(U^{n+1}_i) = 0, \tag{3.7}$$

with

$$\sigma = \frac{\Delta t}{\Delta x},$$

and where the numerical fluxes are given by the *kinetic* flux splitting

$$F^n_{i+\frac{1}{2}} = \mathcal{F}(U^n_i, U^n_{i+1}), \tag{3.8}$$

$$\mathcal{F}(U,V) = F^+(U) + F^-(V), \tag{3.9}$$

$$F^+(U_i) = \int_{\xi\geq 0} \xi \begin{pmatrix} 1 \\ \xi \end{pmatrix} M_i(h, \xi - u)\ d\xi, \tag{3.10}$$

$$F^-(U_i) = \int_{\xi\leq 0} \xi \begin{pmatrix} 1 \\ \xi \end{pmatrix} M_i(h, \xi - u)\ d\xi, \tag{3.11}$$

$$S(U^{n+1}_i) = \begin{pmatrix} 0 \\ g\ h^{n+1}_i\ Z'_i \end{pmatrix}. \tag{3.12}$$

Henceforth, we have obtained a consistent scheme because of the property, again deduced from (2.11),

$$F^+(U) + F^-(U) = \int_{\mathbb{R}} \xi \begin{pmatrix} 1 \\ \xi \end{pmatrix} M(h, \xi - u)\ d\xi = F(U).$$

The conservation of mass and momentum for $Z' = 0$ is also obvious.

Remark 3.1 *The interpretation is that, as usual, the collision term Q, which relaxes f to a Gibbs equilibrium M is neglected in the advection scheme (3.5). And at each timestep we replace $f^{n+1}_i(\xi)$ by $M^{n+1}_i(\xi)$ which is a way to perform all collisions at once and to recover the Gibbs equilibrium without computing it.*

3.3 Stability of the standard kinetic scheme

We now establish the stability property of the kinetic scheme.

Theorem 3.1 *Assume that additionallly, the function χ in (2.6) has a compact support, namely*

$$\exists w_M \in \mathbb{R},\ \textit{such that}\quad \chi(w) = 0\ \textit{for}\ |w| \geq w_M. \tag{3.13}$$

Then, under the CFL condition (independent of Z')

$$\Delta t\ \max(|u^n_i| + w_M \tilde{c}^n_i) \leq \Delta x, \tag{3.14}$$

the kinetic scheme (3.7)–(3.12) keeps the water height positive i.e. $h^n_i \geq 0$ if it is true initially.

Notice that the source term hZ' has no influence on the CFL condition because it is treated implicitely.

Proof of Theorem 3.1. Suppose we have $h_i^n \geq 0 \ \forall i \in \mathbb{Z}$. From the definition of the function M in (2.6) and the positivity of the function χ, we immediately deduce

$$M_i^n(\xi) \geq 0 \qquad \forall i.$$

We now introduce the quantities ξ_+, ξ_- defined by

$$\xi_+ = \max(0, \xi), \quad \xi_- = \max(0, -\xi),$$

and so we rewrite the microscopic scheme (3.5) in the form

$$f_i^{n+1}(\xi - g\Delta t Z_i') = (1 - \sigma|\xi|)M_i^n(\xi) + \sigma\xi_+ M_{i-1}^n(\xi) + \sigma\xi_- M_{i+1}^n(\xi). \tag{3.15}$$

Then, for each j, the value of ξ is either such that

$$|\xi - u_j^n| \geq w_M \tilde{c}_j^n$$

and then from the condition (3.13) on χ and the definition of function M we have

$$M_j^n(\xi) = 0,$$

or the value of ξ is such that

$$|\xi - u_j^n| \leq w_M \tilde{c}_j^n$$

which implies that $|\xi| \leq (|u_j^n| + w_M \tilde{c}_j^n)$ and then, using the CFL condition (3.14), we obtain

$$\sigma|\xi| \leq 1.$$

Therefore, in the relation (3.15), $f_i^{n+1}(\xi - g\Delta t Z_i')$ is a convex combination of nonegative quantities and thus

$$f_i^{n+1}(\xi - g\Delta t Z_i') \geq 0.$$

With a simple integration in ξ, we obtain

$$h_i^{n+1} \geq 0.$$

3.4 A discrete entropy inequality (Z=0)

Another and deeper stability condition is the in-cell entropy inequality i.e. the discrete analog of the inequality (2.2). This is possible for flat bottoms with an appropriate choice for the function χ .

Theorem 3.2 *Consider the particular function χ defined by*

$$\chi(w) = \alpha(1 - \frac{w^2}{\beta})_+^{\frac{1}{2}} \tag{3.16}$$

with $\alpha = 1/\pi$ and $\beta = 4$ in order to satisfy the conditions (2.5). Assume $Z = 0$ and that the CFL condition (3.14) holds. Then the scheme (3.7)–(3.12) satisfies the in-cell entropy inequality

$$E_i^{n+1} - E_i^n + \sigma\ [\eta_{i+\frac{1}{2}}^n - \eta_{i-\frac{1}{2}}^n] \leq 0, \tag{3.17}$$

with the discrete energy and entropy fluxes

$$E_i^n = h_i^n\ \frac{(u_i^n)^2}{2} + \frac{g(h_i^n)^2}{2},$$

$$\eta_{i+\frac{1}{2}}^n = \eta_+(U_i) + \eta_-(U_{i+1}),$$

$$\eta_+(U) = \int_{\xi \geq 0} [\xi^3\ M(h, \xi - u) + \xi\ M(h, \xi - u)^3]\ d\xi,$$

$$\eta_-(U) = \int_{\xi \leq 0} [\xi^3\ M(h, \xi - u) + \xi\ M(h, \xi - u)^3]\ d\xi,$$

Before proving this Theorem, we need an interpretation of the function M as a Gibbs equilibrium.

Lemma 3.1 *The function $M(h, \xi - u)$ built with the shape function χ in (3.16) is the minimizer of the energy*

$$E(f) = \int_{\mathbb{R}} [\frac{\xi^2}{2} f(\xi) + \frac{\pi^2}{6} f(\xi)^3]\ d\xi, \tag{3.18}$$

under the constraints

$$f \geq 0, \quad \int_{\mathbb{R}} f(\xi)\ d\xi = h, \quad \int_{\mathbb{R}} \xi f(\xi)\ d\xi = h\ u.$$

Moreover, recalling the formula (2.1) and (2.11), the minimum is

$$E\Big(M(h,\xi-u)\Big)= E(h,u),$$

where we omit Z which vanishes here.

Proof of Lemma 3.1. The proof of this Lemma is standard since $E(f)$ is a convex functional. We refer to [22] for a proof in the context of full gas dynamics. Notice however that the formula for M (and thus for χ) follows directly from the Euler-Lagrange equation associated to this minimization problem

$$\frac{\xi^2}{2}+\frac{\pi^2}{2}f^2(\xi)=\lambda+\mu\ \xi,$$

where $\lambda(h,u)$ and $\mu(h,u)$ are the Lagrange multipliers. □

Proof of Theorem 3.2. We manipulate the convex relationship (3.5) according to the formula (3.18) for the energy and integrate it in ξ, thanks to the convexity property stated in the proof of the Theorem 3.1, we obtain

$$E(f_i^{n+1})-E_i^n+\frac{\Delta t}{\Delta x}\ [\eta^n_{i+\frac12}-\eta^n_{i-\frac12}]\le 0$$

where the flux has the expression stated in the theorem. Then, we use the Lemma 3.1 to deduce

$$E_i^{n+1}=E(M_i^{n+1})\le E(f_i^{n+1})$$

and this proves the result. □

Notice that the introduction of the source terms does not lead to a conservative entropy inequality as it should be from (2.2). Therefore we loose an important property which, together with nonegativity of water height, gives quadratic a priori bounds on the solution. Nevertheless, an equivalent inequality might occur because, examining the Z terms in the entropy inequality we obtain the approximation

$$(\frac{\partial}{\partial x}hZu)_i^n\approx Z_i'(hu)_i^{n+1}+Z_i[(hu)^n_{i+\frac12}-(hu)^n_{i-\frac12}]+h_i^{n+1}(Z_i'\ \Delta t)^2.$$

Remark 3.2 *In order to simplify the actual implementation of the kinetic scheme, we use the particular function χ in (2.13) rather than the more complex choice which leads to the entropy inequality. It seems, even though it has not been proved, that the scheme determined with (2.13) is entropic.*

4 1D numerical tests

We consider here two simple numerical test cases for which an analytical solution can be computed. Analogous problems have been proposed in a Workshop initiated by Electricité de France (see ([13]) for the proceedings of this workshop). With the second test case we show the loss of efficiency of the standard scheme (3.7)–(3.12) in some situations of equilibrium with variable bottom.

4.1 Dam-break

We begin with a non-stationary test case. It is a simple dam-break in a channel with flat bottom ($Z = 0$). In this case the entropy theory (see Sec.3.4) is valid. The initial conditions are

$$h(0,x) = \begin{cases} h_l, & \text{for } x \leq 0, \\ h_r, & \text{for } x > 0, \end{cases} \tag{4.1}$$

$$u(0,x) = 0. \tag{4.2}$$

We choose $h_l > h_r$ to be coherent with the physical phenomenon of a dam-break from the left to the right. This test allows to isolate the behaviour of the scheme on a shock problem. See ,e.g., ([8], [27], [28]) for the computation of the exact solution.

The computational domain is symmetric around the point $x = 0$ and the total length L is 2,000 meters. We have used two different values for the space step, $L/100$ and $L/1000$ in order to show the effects of the mesh size. The timestep is computed thanks to the CFL condition (3.14). We have tested the dam-break on a wet bed and on a dry soil.

For the first test case, we prescribe $h_l = 1m$ and $h_r = 0.5m$, the Figure 4.1 (resp 4.2 present the water level (resp. the discharge) for $\Delta x = L/100$ at time $t = 200s$. In Figures 4.3, 4.4 we present the same results for $\Delta x = L/1000$. We compare the computed solutions plotted with a continuous line with the exact one (dot lines). The position of the shock is in good agreement with the exact solution, but on the coarse mesh the shock is smoothed as usual with a first order scheme. For the second test case, we consider a dry bed on the right part of the channel, the computed water level and discharge are also compared at time $t = 150s$ to the exact ones in Figures 4.5-4.8. We verify numerically that the water height positivity is preserved.

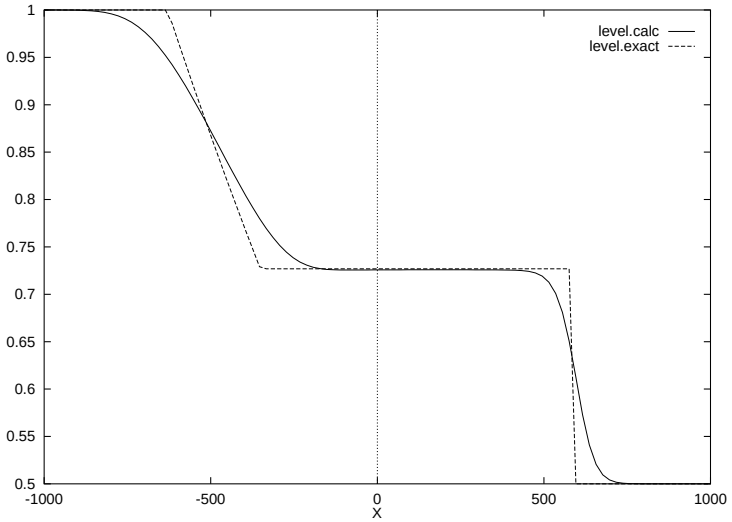

Figure 4.1: Dam break: Water level, $\Delta x = L/100$.

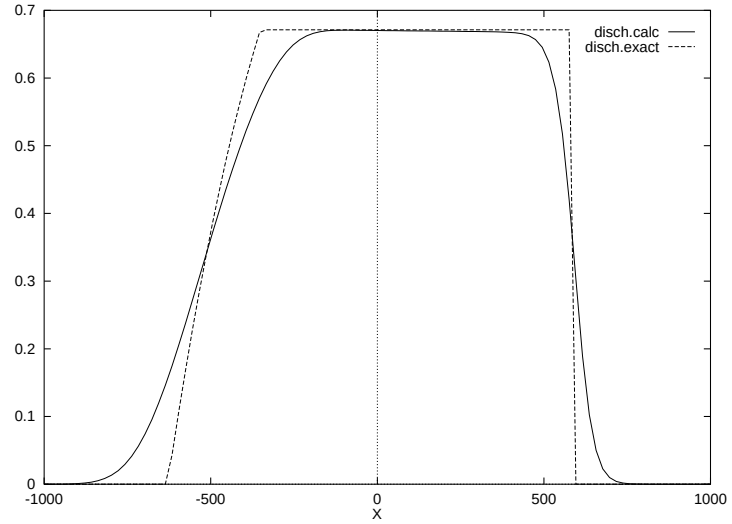

Figure 4.2: Dam break: Discharge, $\Delta x = L/100$.

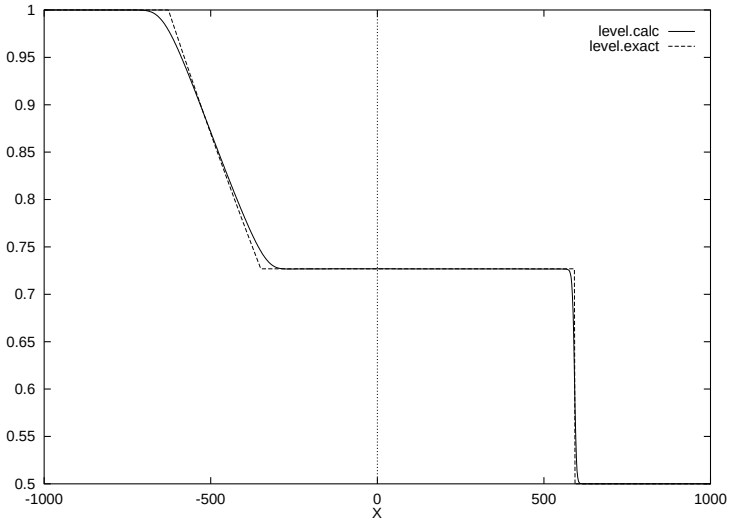

Figure 4.3: Dam break: Water level, $\Delta x = L/1000$.

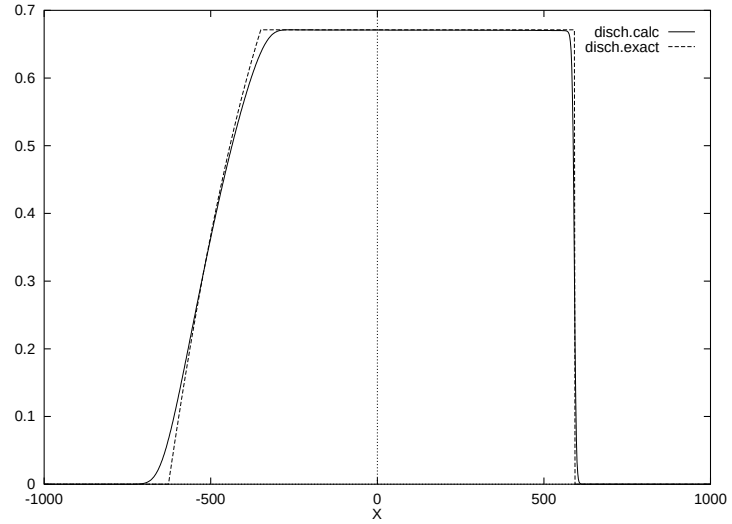

Figure 4.4: Dam break: Discharge, $\Delta x = L/1000$.

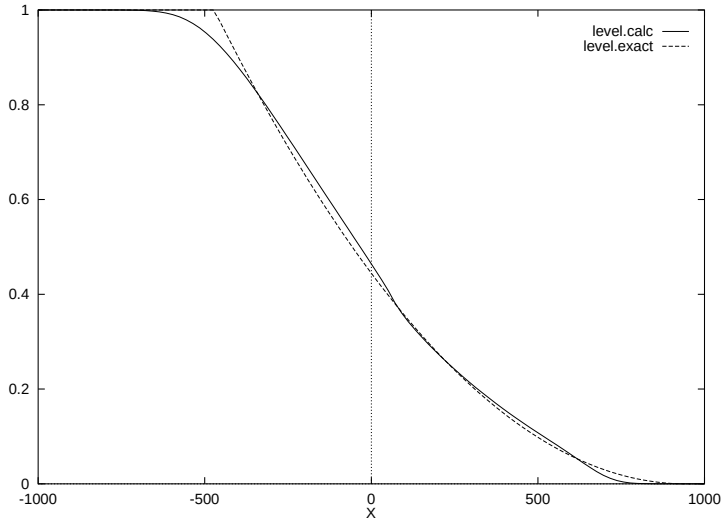

Figure 4.5: Dam break on a dry bed: Water level, $\Delta x = L/100$.

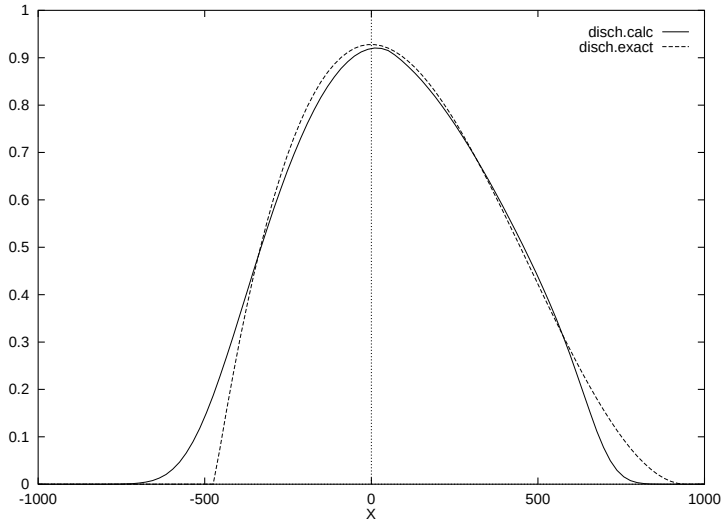

Figure 4.6: Dam break on a dry bed: Discharge, $\Delta x = L/100$.

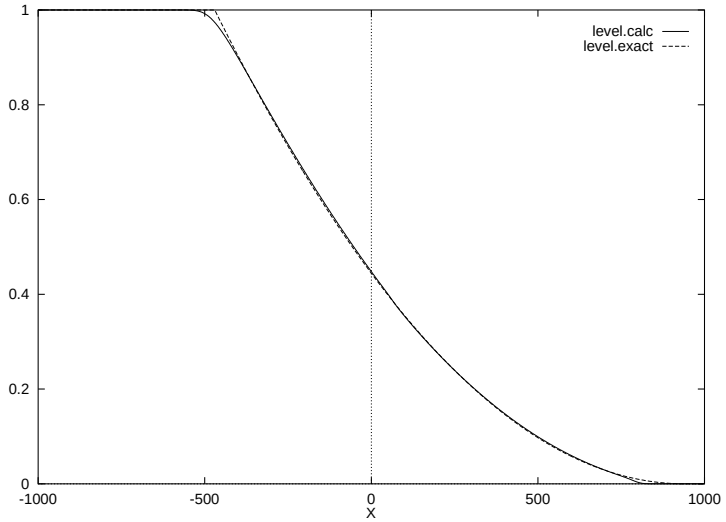

Figure 4.7: Dam break on a dry bed: Water level, $\Delta x = L/1000$.

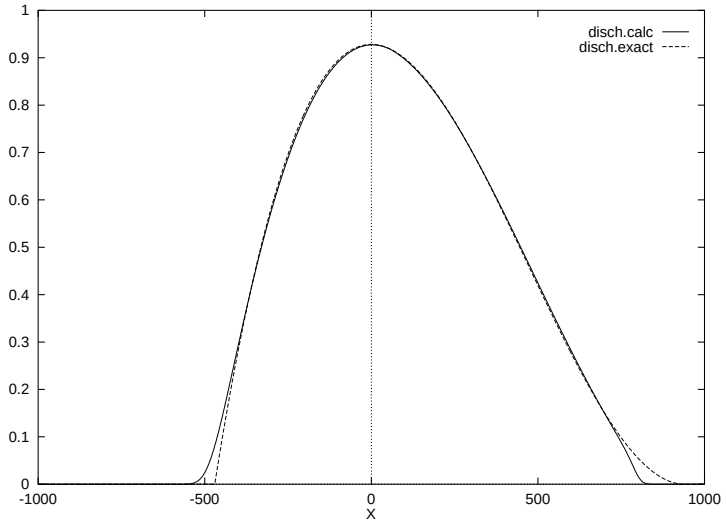

Figure 4.8: Dam break on a dry bed: Discharge, $\Delta x = L/1000$.

4.2 Still Water

We now study a test case concerning the still water equilibrium. The computational domain is a channel with a parabolic bump on the bottom modelized by the equation

$$Z(x) = [0.2 - 0.05(x - 10)^2]^+. \tag{4.3}$$

The length of the channel is 20m, we use $\Delta x = L/100$. The initial conditions are

$$\begin{aligned} u(t,x) &= 0,, \\ h(t,x) + Z(x) &= 2. \end{aligned}$$

This steady state is not preserved by the standard kinetic scheme as shown in Figures (4.9), (4.10) which present the solution at time $t = 50s$. Some oscillations appear on the free surface and the velocity field near the bump is not 0.

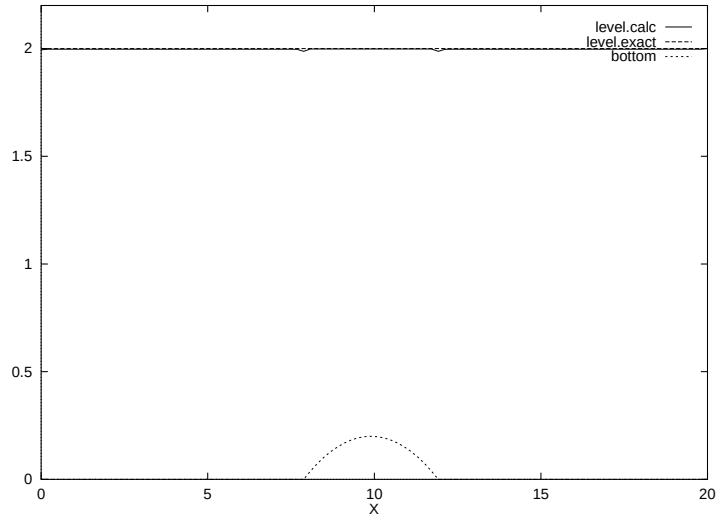

Figure 4.9: Still water: Water level, standard kinetic scheme.

5 A well-balanced kinetic scheme

We now present a variant of the scheme which is able to preserve the still water equilibrium and keep the stability properties of kinetic schemes (height positivity

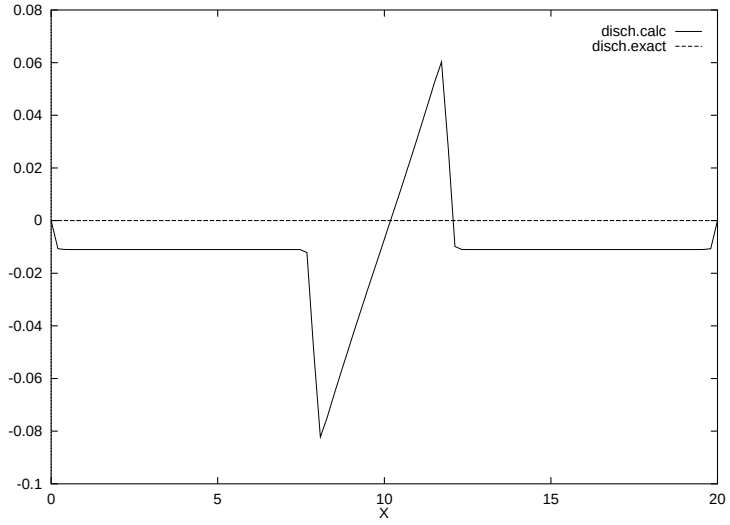

Figure 4.10: Still water: Discharge, standard kinetic scheme.

and entropy inequality). Another extension of kinetic schemes which preserves the still water equilibrium will be presented in a forthcoming paper ([24]).

5.1 General form of the scheme

The flux splitting form of the 'characteristic' scheme which takes into account the bottom Z is incompatible with the equilibrium conservation property . The fluxes have to take into account somehow the neighboring cells. Hence we introduce two quantities h_i^+ and h_i^- that will be precised later, but always have the following dependency

$$h_i^+ = \mathcal{H}(h_i, Z_i, h_{i+1}, Z_{i+1}),$$
$$h_i^- = \mathcal{H}(h_i, Z_i, h_{i-1}, Z_{i-1}).$$

the function $\mathcal{H}$ is chosen so as to ensure the following constraints

$$\exists \alpha > 0 \text{ such that } \forall i \in \mathbb{Z}, \qquad 0 \leq h_i^{\pm} \leq \alpha h_i, \tag{5.1}$$

$$h_i + Z_i = h_{i+1} + Z_{i+1} \Longrightarrow h_i^+ = h_{i+1}^-. \tag{5.2}$$

We use the expression of the fluxes $F^+(U_i)$ and $F^-(U_i)$ introduced in (3.10) and (3.11) and the definition of M in (2.6) to write

$$\begin{aligned} F^+(U_i) &= h_i \int_{w \geq -\frac{u_i}{\tilde{c}_i}} (u_i + w\tilde{c}_i) \begin{pmatrix} 1 \\ u_i + w\tilde{c}_i \end{pmatrix} \chi(w) dw, \\ F^-(U_i) &= h_i \int_{w \leq -\frac{u_i}{\tilde{c}_i}} (u_i + w\tilde{c}_i) \begin{pmatrix} 1 \\ u_i + w\tilde{c}_i \end{pmatrix} \chi(w) dw \end{aligned}$$

and we introduce the corrected fluxes (we use the same notation)

$$\begin{aligned} F^+(U_i) &= h_i \int_{w \geq -\frac{u_i h_i}{h_i^+ \tilde{c}_i^+}} (u_i + w\tilde{c}_i^+ \tfrac{h_i^+}{h_i}) \begin{pmatrix} 1 \\ u_i + w\tilde{c}_i^+ \frac{h_i^+}{h_i} \end{pmatrix} \chi(w) dw, \\ F^-(U_i) &= h_i \int_{w \leq -\frac{u_i h_i}{h_i^- \tilde{c}_i^-}} (u_i + w\tilde{c}_i^- \tfrac{h_i^-}{h_i}) \begin{pmatrix} 1 \\ u_i + w\tilde{c}_i^- \frac{h_i^-}{h_i} \end{pmatrix} \chi(w) dw. \end{aligned} \tag{5.3}$$

We recall the notation $\tilde{c}_i^{\pm} = \sqrt{\frac{g h_i^{\pm}}{2}}$. At this level the source term is taken explicit, so in (3.7) it is denoted $S(U_i^n)$, it is not specified at this stage except that its first component is still assumed to vanish.

Thanks to these transformations we obtain a new scheme that has some theoretical properties that we present in the next subsection. In order to simplify the notations we will drop the superscript n when no mistake is possible.

5.2 Properties of the scheme

Theorem 5.1 *If $h_i^{\pm} = h_i$ for $h_i = h_{i\pm1}$, and $Z_i = Z_{i\pm1}$, then the kinetic scheme (3.7)-(3.9), (5.3) is consistent.*

Proof of the Theorem 5.1. We have to check that $F^+(U) + F^-(U) = F(U)$ which follows directly from the assumptions of the theorem. □

Theorem 5.2 *We assume (5.1) and that the function χ satisfies the support condition (3.13) then under the CFL condition on the time step*

$$\Delta t \ \max(|u_i^n| + \alpha^{\frac{3}{2}} w_M \tilde{c}_i^n) \leq \Delta x, \tag{5.4}$$

the kinetic scheme (3.7)-(3.9), (5.3) preserves the water height positivity.

Proof of the Theorem 5.2. We give the proof directly at the macroscopic level. By induction, we have

$$\forall i \in \mathbb{Z} \qquad h_i^n \geq 0.$$

From (3.7)–(3.9) we have

$$h_i^{n+1} = h_i^n + \sigma[F_h^+(U_{i-1}^n) + F_h^-(U_i^n) - F_h^+(U_i^n) - F_h^-(U_{i+1}^n)]$$

and as we know that

$$\forall U \qquad F_h^+(U) \geq 0, \qquad F_h^-(U) \leq 0,$$

it remains to prove that

$$I = h_i^n + \sigma[F_h^-(U_i^n) - F_h^+(U_i^n)] \geq 0.$$

Considering the fluxes defined in (5.3), we see that

$$\begin{aligned} I = h_i^n \Big(1 \quad &+ \quad \sigma \int_{w \leq -\frac{u_i h_i}{h_i^- \tilde{c}_i^-}} (u_i + w\tilde{c}_i^- \frac{h_i^-}{h_i}) \chi(w) dw \\ &- \quad \sigma \int_{w \geq -\frac{u_i h_i}{h_i^+ \tilde{c}_i^+}} (u_i + w\tilde{c}_i^+ \frac{h_i^+}{h_i}) \chi(w) dw \Big). \end{aligned}$$

We use the relationship

$$\int_{w \leq -\frac{u_i h_i}{h_i^- \tilde{c}_i^-}} u_i \chi(w) dw - \int_{w \geq -\frac{u_i h_i}{h_i^+ \tilde{c}_i^+}} u_i \chi(w) dw \;\geq\; -|u_i|,$$

and the following consequence of the support property (3.13)

$$\begin{aligned} \int_{w \leq -\frac{u_i h_i}{h_i^- \tilde{c}_i^-}} w\tilde{c}_i^- \frac{h_i^-}{h_i} \chi(w) dw \;&\geq\; -\frac{w_M}{2} \tilde{c}_i^- \frac{h_i^-}{h_i}, \\ \int_{w \leq -\frac{u_i h_i}{h_i^+ \tilde{c}_i^+}} w\tilde{c}_i^+ \frac{h_i^+}{h_i} \chi(w) dw \;&\leq\; \frac{w_M}{2} \tilde{c}_i^+ \frac{h_i^+}{h_i}. \end{aligned}$$

to obtain finally

$$I \geq h_i^n [1 - \sigma(|u_i| + \frac{w_M}{2}(\tilde{c}_i^- \frac{h_i^-}{h_i} + \tilde{c}_i^+ \frac{h_i^+}{h_i}))].$$

Using the upper bound on h_i^- and h_i^+, and the definition of $\tilde{c}$ we conlude that

$$I \geq h_i^n[1 - \sigma(|u_i| + w_M \alpha^{\frac{3}{2}} \tilde{c}_i)].$$

Thanks to the CFL condition, this proves the positivity of I and concludes the proof. □

5.3 Equilibriums

We now prove that it is possible to preserve the still water equilibrium. We consider first the water height equation, then the momentum equation .

Lemma 5.1 *We assume that the property (5.2) is satisfied, then if the solution at step n verifies*

$$u_i^n = 0, \quad h_i^n + Z_i = H, \quad \forall i \in \mathbb{Z}$$

the scheme (3.7)–(3.9), (5.3) ensures that

$$h_i^{n+1} = h_i^n, \quad \forall i \in \mathbb{Z}.$$

Proof of Lemma 5.1. A direct calculation of the first component of the fluxes shows easily that $F_h^+(U_i) + F_h^-(U_{i+1}) = 0$, thus proving the Lemma. □

We now try to find an appropriate discretisation of the source term in order to ensure that the second equation also preserves the still water equilibrium.

Theorem 5.3 *We assume again (5.2), and choose the source term*

$$S(U_i^n) = \begin{pmatrix} 0 \\ g\tilde{h}_i \frac{Z_{i+\frac{1}{2}} - Z_{i-\frac{1}{2}}}{\Delta x} \end{pmatrix}, \tag{5.5}$$

with

$$\tilde{h}_i = \frac{h_{i-\frac{1}{2}} + h_{i+\frac{1}{2}}}{2},$$

where $h_{i-\frac{1}{2}}$ and $h_{i-\frac{1}{2}}$ will be precised later and where the quantities $Z_{i+\frac{1}{2}}$ and $Z_{i-\frac{1}{2}}$ have the following properties

$$h_i + Z_i = h_{i\pm1} + Z_{i\pm1} \Longrightarrow h_{i-\frac{1}{2}} + Z_{i-\frac{1}{2}} = h_{i+\frac{1}{2}} + Z_{i+\frac{1}{2}}, \tag{5.6}$$

$$Z_i = Z_{i+1} \Longrightarrow Z_{i+\frac{1}{2}} = Z_i = Z_{i+1}. \tag{5.7}$$

Then the scheme preserves the still water equilibrium.

Proof of Theorem 5.3. From (3.7)–(3.9), (5.3), one can compute

$$\begin{aligned} q_i^{n+1} &= q_i^n + \sigma[F_q^+(U_{i-1}^n) + F_q^-(U_i^n) - F_q^+(U_i^n) - F_q^-(U_{i+1}^n)] - \Delta t S_q(U_i^n) \\ &= q_i^n + \frac{\sigma}{2}[(\frac{{h_{i-1}^+}^2 \tilde{c}_{i-1}^{+\,2}}{h_{i-1}} + \frac{h_i^{-\,2}\tilde{c}_i^{-\,2}}{h_i}) - (\frac{h_i^{+2}\tilde{c}_i^{+2}}{h_i} + \frac{{h_{i+1}^-}^2 \tilde{c}_{i+1}^{-\,2}}{h_{i+1}})] - \Delta t S_q(U_i^n). \end{aligned}$$

Then, we choose

$$h_{i+\frac{1}{2}} = \frac{1}{\sqrt{g}}\sqrt{\frac{h_i^{+\,2}\tilde{c}_i^{+2}}{h_i} + \frac{{h_{i+1}^-}^2 \tilde{c}_{i+1}^{-\,2}}{h_{i+1}}},$$

and we obtain

$$q_i^{n+1} = q_i^n - g\sigma\tilde{h}_i(h_{i+\frac{1}{2}} - h_{i-\frac{1}{2}}) - \Delta t S_q(U_i^n).$$

Using the definition of $S(U_i^n)$ in (5.5), we find

$$q_i^{n+1} = q_i^n - g\sigma\tilde{h}_i[(h_{i+\frac{1}{2}} + Z_{i+\frac{1}{2}}) - (h_{i-\frac{1}{2}} + Z_{i-\frac{1}{2}})],$$

and the result is proved. □

5.4 Choice of the interpolations

A simple choice satisfying the set of properties (5.1), (5.2), (5.6), (5.7) is

$$\begin{aligned} h_i^+ &= h_i + Z_i - \min(\max(Z_i, Z_{i+1}), h_i + Z_i), \\ h_i^- &= h_i + Z_i - \min(\max(Z_i, Z_{i-1}), h_i + Z_i), \\ Z_{i+\frac{1}{2}} &= \sqrt{\frac{(h_i + Z_i)^2 + (h_{i+1} + Z_{i+1})^2}{2}} \; - \; h_{i+\frac{1}{2}}. \end{aligned}$$

The condition (5.1) is satisfied with $\alpha = 1$ so the CFL condition (see (3.14),(5.4)) is not modified. Moreover with these formulae we recover the standard scheme for flat bottoms.

5.5 Numerical results

We have tested this new scheme on the still water problem defined in Sec 4.2., then the equilibrium is preserved up to the machine accuracy as shown in Figure (5.1) where the error on the water level is plotted and in Figure (5.2) where the discharge is plotted.

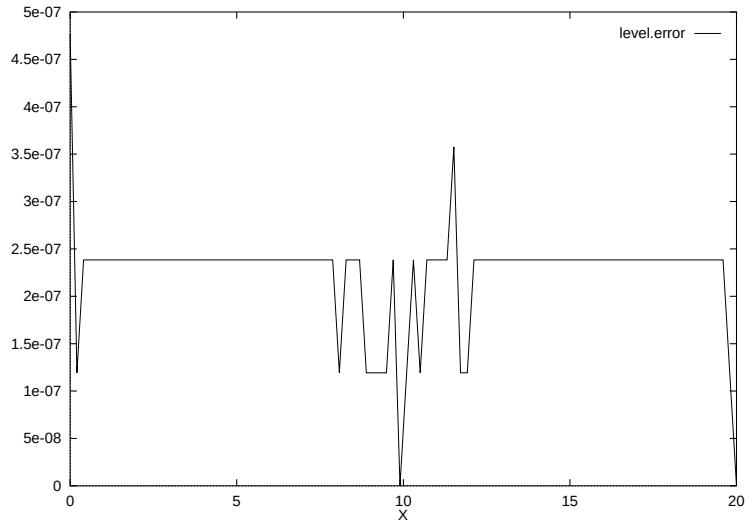

Figure 5.1: Still water: Water level error, equilibrium scheme.

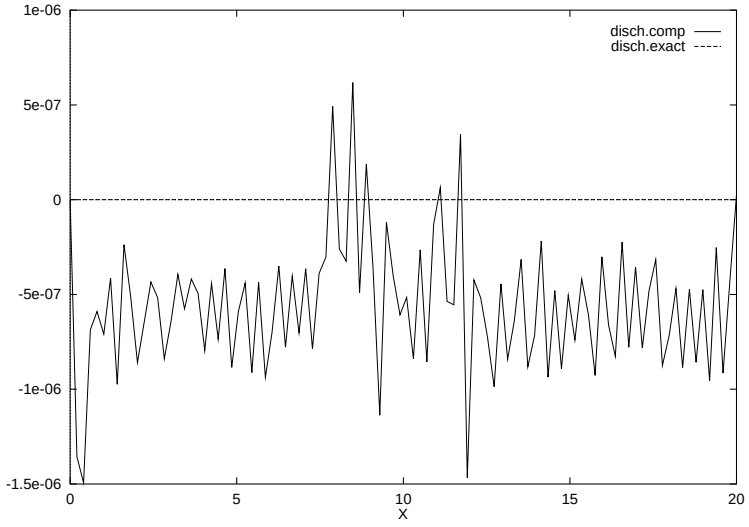

Figure 5.2: Still water: Discharge, equilibrium scheme.

The following test cases concern the same geometry with a constant discharge prescribed on the inflow boundary, then the conditions (2.4) are verified and we can deduce the exact water height. Three cases are considered, dealing with a subcritical (or fluvial) flow, a transcritical one without shock (the flow becomes torrential on the bump and the outflow is torrential) and a transcritical one with shock (the flow becomes torrential on the bump and the outflow is fluvial). These test cases have also been treated in the Workshop initiated by Electricité de France ([13]). The different cases are defined by the boundary conditions, we prescribe the inflow discharge q_{in} and the outflow water level H_{out} except for the torrential outflow:

- Fluvial flow: $q_{in} = 4.42\ m^3/s, \quad H_{out} = 2.m$
- Transcritical flow without shock : $q_{in} = 1.53\ m^3/s, \quad H_{out} = 0.66m$
- Transcritical flow with shock : $q_{in} = 0.18\ m^3/s, \quad H_{out} = 0.33m$

We use $\Delta x = L/100$. The initial conditions are given by constant data equal to the boundary conditions. For each case we have plotted the computed water level and the discharge compared to the exact ones in Figures (6.1)-(6.6). The constant discharge is not preserved, the error is of the order of 1% except in the shock where it is larger.

6 Adapted kinetic scheme for 2D Saint-Venant equations

We now want to apply our new scheme to two-dimensional problems. We begin with the presentation of the finite volume method on a general triangular grid and some sentences about the general kinetic theory in 2D. Then in subsections 6.3 and 6.4 we explain the 2D kinetic scheme and give some details on the numerical implementation, finally in subsection 6.5 we present the 2D extension of the equilibrium scheme.

6.1 2D Finite Volume method

The general method is close to the 1D finite volume method developed in subsection 3.1.

Let $\mathcal{T}_h$ be a triangulation of $\mathbb{R}^2$ which vertices are denoted P_i. The dual cells C_i are obtained by joining the centers of mass of the triangles surrounding each vertex P_i. We use the following notations (see Fig(6.7)):

K_i, set of nodes P_j surrounding P_i,

$|C_i|$, area of C_i,

Γ_{ij}, boundary edge belonging to cells C_i and C_j,

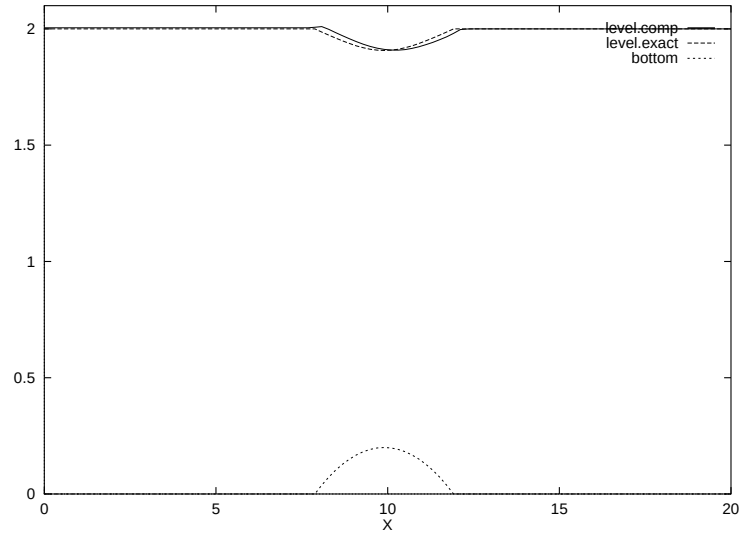

Figure 6.1: Constant discharge, Fluvial flow, Water level

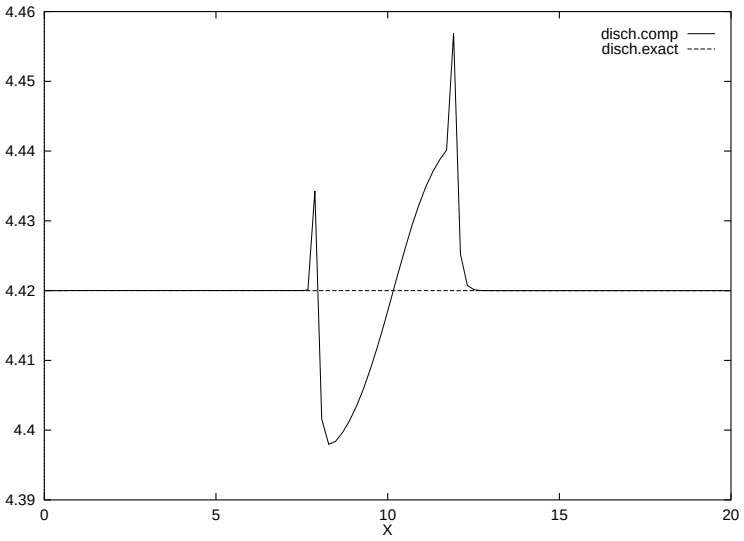

Figure 6.2: Constant discharge, Fluvial flow, Discharge

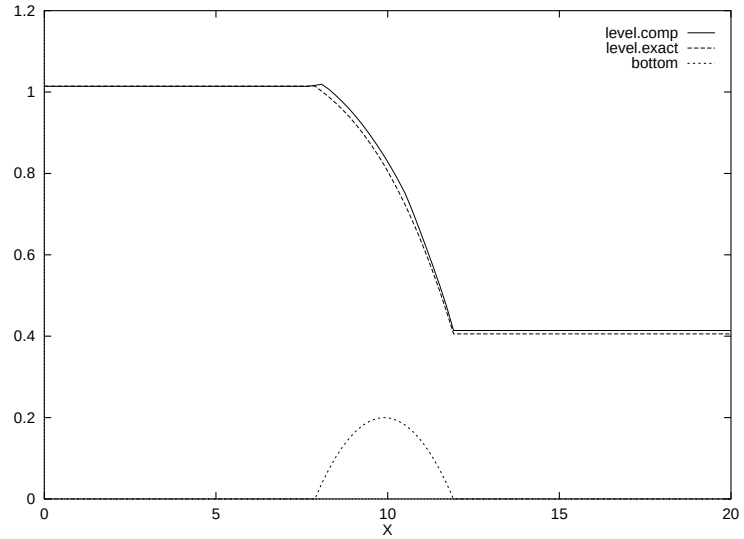

Figure 6.3: Constant discharge, Torrential flow without shock, Water level.

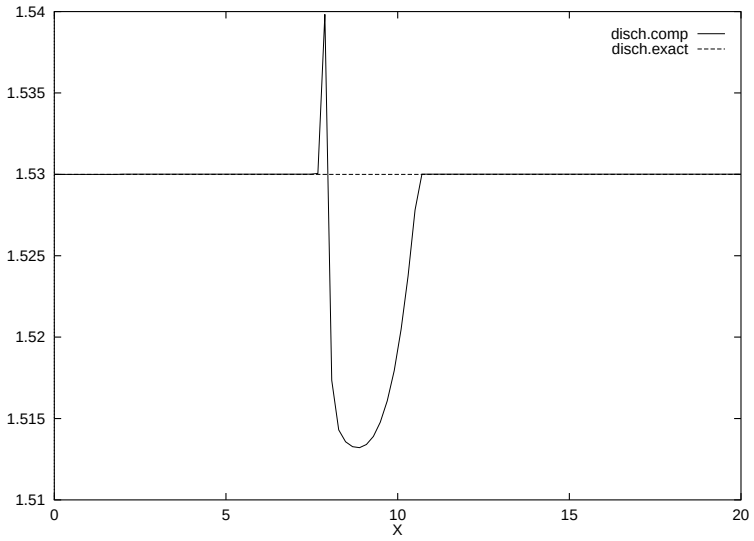

Figure 6.4: Constant discharge, Torrential flow without shock, Discharge.

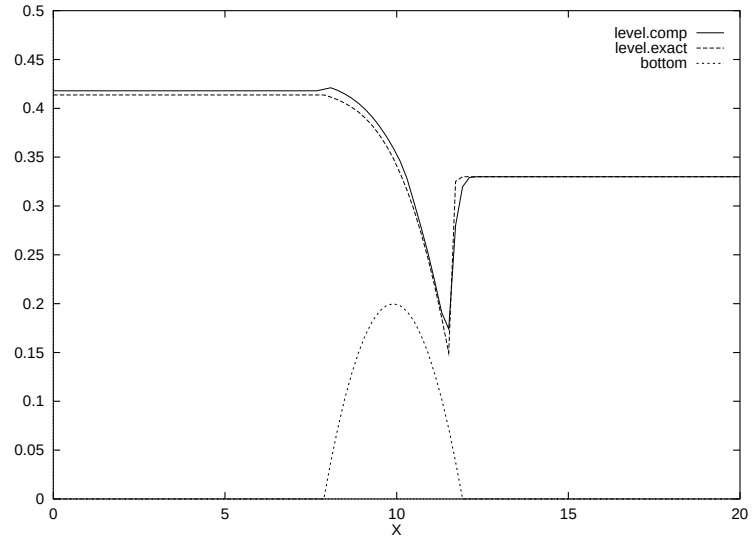

Figure 6.5: Constant discharge, Torrential flow with shock, Water level.

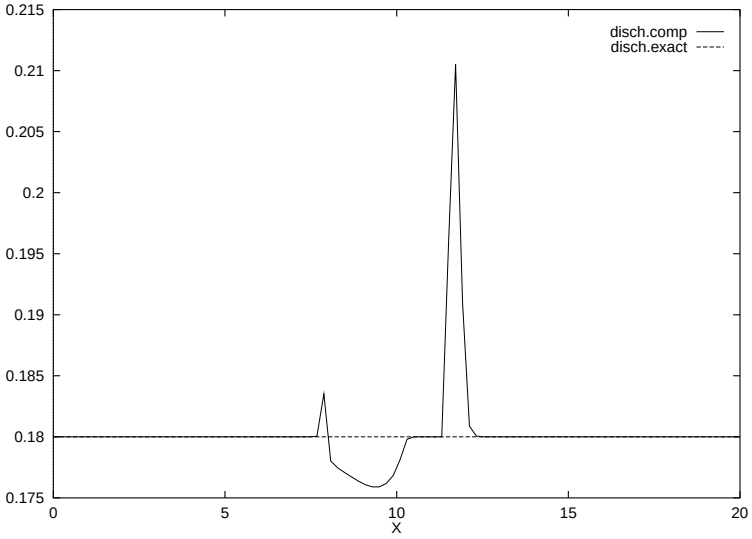

Figure 6.6: Constant discharge, Torrential flow with shock, Discharge.

L_{ij}, length of Γ_{ij},
n_{ij}, unit normal to Γ_{ij}, outward to C_i (n_{ji}=-n_{ij}).

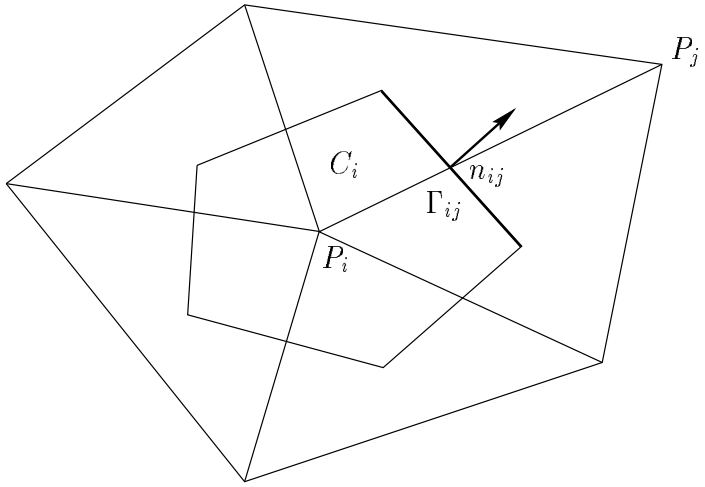

Figure 6.7: Dual cell C_i.

We always consider a system of first order conservation laws

$$\frac{\partial U}{\partial t} + \mathrm{div} F(U) = 0. \tag{6.1}$$

For Saint-Venant system (1.1),(1.2) with $\nabla Z = 0$, we have:

$$U = \begin{pmatrix} h \\ q \end{pmatrix}, \qquad F(U) = \begin{pmatrix} q \\ \dfrac{q \otimes q}{h} + \dfrac{g}{2} h^2 Id \end{pmatrix}. \tag{6.2}$$

Then the finite volume scheme writes

$$U_i^{n+1} = U_i^n - \sum_{j \in K_i} \sigma_{ij} \mathcal{F}(U_i^n, U_j^n, n_{ij}) \tag{6.3}$$

with

$$\sigma_{ij} = \frac{\Delta t \, L_{ij}}{|C_i|}, \tag{6.4}$$

where $\mathcal{F}(U_i, U_j, n_{ij})$ denotes an interpolation of the normal component of the flux $F(U) \cdot n_{ij}$ along the edge Γ_{ij}. This interpolation is usually performed using a one dimensional solver since locally the problem looks like a planar discontinuity.

6.2 2D Kinetic formulation

As in Sec.2.2, we introduce $\chi(w)$, a positive, even function defined on $\mathbb{R}^2$ and satisfying

$$\int_{\mathbb{R}^2} \begin{pmatrix} 1 \\ w_i w_j \end{pmatrix} \chi(w)\ dw = \begin{pmatrix} 1 \\ \delta_{ij} \end{pmatrix}. \tag{6.5}$$

In addition we assume that $\chi(w)$ is compactly supported, i.e.

$$\exists w_M \in \mathbb{R}, \text{ such that } \quad \chi(w) = 0 \text{ for } |w| \geq w_M. \tag{6.6}$$

An example of function χ satisfying these properties is

$$\chi(w) = \frac{1}{12} \mathbb{1}_{|w_i| \leq \sqrt{3}},\ i = 1, 2. \tag{6.7}$$

In 2D, the density of particles $M(t, x, \xi)$ is written

$$M(t, x, \xi) = \frac{h(t, x)}{\tilde{c}^2}\ \chi(\frac{\xi - u(t, x)}{\tilde{c}}) \tag{6.8}$$

with $\tilde{c}$ defined by (2.7).

We have a 2D theorem equivalent to Theorem 2.2.

6.3 2D Kinetic scheme

We follows here Sec.3.2. We always assume in this section $\nabla Z = 0$. The discretization of the microscopic equation by a simple upwind scheme gives

$$f_i^{n+1}(\xi) - M_i^n(\xi) + \frac{\Delta t}{|C_i|} \sum_{j \in K_i} L_{ij}\ \xi \cdot n_{ij}\ M_{ij}^n(\xi) = 0, \tag{6.9}$$

with the fluxes $M_{ij}^n(\xi)$ computed by the upwind formula

$$M_{ij}^n(\xi) = \begin{cases} M_i^n(\xi) & \text{for } \xi \cdot n_{ij} \geq 0, \\ M_j^n(\xi) & \text{for } \xi \cdot n_{ij} \leq 0. \end{cases}$$

Then we compute U_i^{n+1} by

$$U_i^{n+1} = \int_{\mathbb{R}^2} \begin{pmatrix} 1 \\ \xi \end{pmatrix} f_i^{n+1}(\xi)\ d\xi. \tag{6.10}$$

The equation (6.10) can be written with the form (6.3) with

$$\mathcal{F}(U_i, U_j, n_{ij}) = F^+(U_i, n_{ij}) + F^-(U_j, n_{ij}), \tag{6.11}$$

and

$$F^+(U_i, n_{ij}) = \int_{\xi \cdot n_{ij} \geq 0} \xi \cdot n_{ij} \begin{pmatrix} 1 \\ \xi \end{pmatrix} M_i(\xi)\, d\xi, \tag{6.12}$$

$$F^-(U_j, n_{ij}) = \int_{\xi \cdot n_{ij} \leq 0} \xi \cdot n_{ij} \begin{pmatrix} 1 \\ \xi \end{pmatrix} M_j(\xi)\, d\xi. \tag{6.13}$$

It is clear that the scheme is consistent and conservative. We can also prove that the water height positivity is preserved under the CFL condition

$$\Delta t \, \max(|u_i^n| + w_M \tilde{c}_i^n) \leq \frac{|C_i|}{\sum_{j \in K_i} L_{ij}}. \tag{6.14}$$

6.4 2D Numerical Implementation

We give here some details on the implementation of the kinetic scheme defined by (6.3),(6.11)-(6.13). For each edge Γ_{ij}, we define a local basis (x_1, x_2) associated to the normal direction and to the tangential one. We denote $\hat{U}$, the vector deduced from U by the rotation in this new basis. If $n_{ij} = (n_x, n_y)$, we have $\hat{U}$ defined by

$$\hat{U} = R\, U \quad \text{with} \quad R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & n_x & n_y \\ 0 & -n_y & n_x \end{pmatrix}$$

and

$$F^+(U, n) = R^{-1} \hat{F}^+(\hat{U}), \qquad R^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & n_x & -n_y \\ 0 & n_y & n_x \end{pmatrix}.$$

We give the detailed expression of $\hat{F}^+(\hat{U}_i)$ related to the edge Γ_{ij}

$$\hat{F}^+(\hat{U}_i) = \frac{h_i}{\tilde{c}_i^2} \int_{(\xi_1 \geq 0) \times \mathbb{R}} \xi_1 \begin{pmatrix} 1 \\ \xi \end{pmatrix} \chi\left(\frac{\xi - \hat{u}_i}{\tilde{c}_i}\right) d\xi \tag{6.15}$$

or

$$\hat{F}^+(\hat{U}_i) = h_i \int_{(w_1 \geq \frac{-\hat{u}_{i,1}}{\tilde{c}_i}) \times \mathbb{R}} (\hat{u}_{i,1} + w_1\, \tilde{c}_i) \begin{pmatrix} 1 \\ \hat{u}_i + w\, \tilde{c}_i \end{pmatrix} \chi(w)\, dw, \tag{6.16}$$

we obtain F^- by an analogous computation.

6.5 The 2D equilibrium scheme

For the cases where $\nabla Z \neq 0$ we write the scheme with the form

$$U_i^{n+1} = U_i^n - \sum_{j \in K_i} \sigma_{ij} \mathcal{F}(U_i^n, U_j^n, n_{ij}) - \Delta t S(U_i^n) \tag{6.17}$$

where

$$S(U_i^n) = \begin{pmatrix} 0 \\ S_q(U_i^n) \end{pmatrix} \tag{6.18}$$

with $S_q(U_i^n)$ an approximation of

$$\frac{g}{|C_i|} \int_{C_i} h(t^n, x)\ \nabla Z dx. \tag{6.19}$$

As for the 1D problem, in order to obtain a scheme preserving the still water equilibrium, we define corrected fluxes. We introduce first a modified value of h_i in the fluxes to preserve the constant value of $h_i^n + Z_i$, this extension from 1D to 2D is straightforward. Then we have to define $S(U_i^n)$ in such a way to preserve the condition $u_i^n = 0$, this extension is less obvious and we have to introduce a sub-cell method to be able to define the value equivalent to $\tilde{h}_i$ in (5.5).

In (6.17) we replace the fluxes defined by (6.16) by

$$\hat{F}^+(\hat{U}_i) = h_i \int_{w_1 \geq \frac{-\hat{u}_{i,1} h_i}{\tilde{c}_{ij} h_{ij}}} \left(\hat{u}_{i,1} + w_1 \tilde{c}_{ij} \frac{h_{ij}}{h_i} \right) \begin{pmatrix} 1 \\ \hat{u}_i + w\, \tilde{c}_{ij} \frac{h_{ij}}{h_i} \end{pmatrix} \chi(w)\, dw \tag{6.20}$$

with

$$h_{ij} = h_i + Z_i - \min(\max(Z_i, Z_j), h_i + Z_i). \tag{6.21}$$

We then have a result analogous to Lemma 5.1 and to the properties of Section 5.2.

With these new fluxes, if we write the momentum equation with $u_i^n = 0$, we obtain

$$q_i^{n+1} = q_i^n - \frac{g}{2} \sum_{j \in K_i} \sigma_{ij}\ \bar{h}_{ij}^{n\,2}\ n_{ij} - \Delta t S_q(U_i^n) \tag{6.22}$$

with

$$\bar{h}_{ij}^2 = \frac{h_{ij}^2 \tilde{c}_{ij}^2}{h_i} + \frac{h_{ji}^2 \tilde{c}_{ji}^2}{h_j}. \tag{6.23}$$

We use the property

$$\sum_{j \in K_i} L_{ij} n_{ij} = 0,$$

so we can write

$$q_i^{n+1} = q_i^n - \frac{g}{2} \sum_{j \in K_i} \sigma_{ij} \ (\bar{h}_{ij}^2 - h_i^2) \ n_{ij} - \Delta t S_q(U_i^n). \tag{6.24}$$

If we define

$$\tilde{h}_{ij} = \frac{(\bar{h}_{ij} + h_i)}{2}, \tag{6.25}$$

the equation (6.24) can be written

$$q_i^{n+1} = q_i^n - g \sum_{j \in K_i} \sigma_{ij} \ \tilde{h}_{ij} (\bar{h}_{ij} - h_i) \ n_{ij} - \Delta t S_q(U_i^n). \tag{6.26}$$

For each edge (P_i, P_j) we then have defined a term $\tilde{h}_{ij}$ which is equivalent to the term $\tilde{h}_i$ in (5.5) for the segment $[i-1, i+1]$. In order to define $S_q(U_i^n)$, we consider in each cell C_i the triangles obtained by joining the ends of the edges Γ_{ij} to the node P_i, we denote T_{ij} this triangle (see Fig.(6.8)). We have

$$\frac{1}{|C_i|} \int_{C_i} h(t^n, x) \ \nabla Z dx = \frac{1}{|C_i|} \sum_{j \in K_i} \int_{T_{ij}} h(t^n, x) \nabla Z \ dx. \tag{6.27}$$

We can consider $\tilde{h}_{ij}$ as an average value of h on T_{ij} and using the integration by parts

$$\int_{T_{ij}} \nabla Z \ dx = \int_{\partial T_{ij}} Z \cdot n \ d\Gamma \approx L_{ij} n_{ij} (Z_{ij} - Z_i) \tag{6.28}$$

with Z_{ij} some approximation of Z on Γ_{ij}, we finally define

$$\Delta t S_q(U_i) = g \sum_{j \in K_i} \sigma_{ij} n_{ij} \tilde{h}_{ij} \, (Z_{ij} - Z_i). \tag{6.29}$$

We can enounce a 2D equilibrium theorem

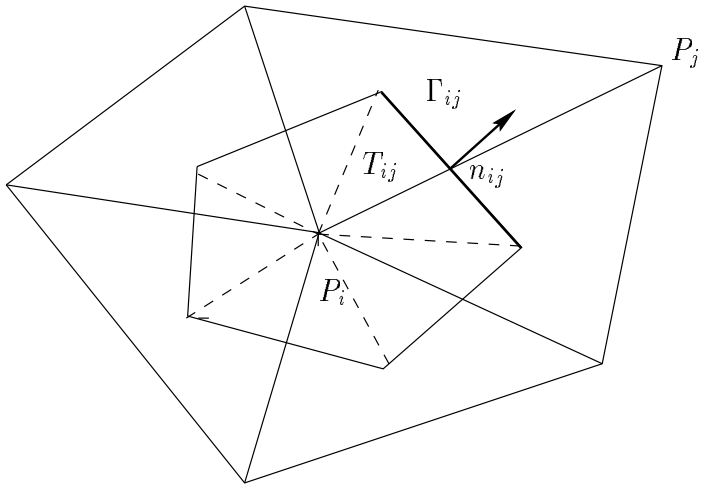

Figure 6.8: Subcells T_{ij}.

Theorem 6.1 *If the quantities Z_{ij} introduced in the discretization of the source term (6.18), (6.29) are such that*

$$\forall i,j \qquad h_i + Z_i = h_j + Z_j \Longrightarrow \bar{h}_{ij} + Z_{ij} = h_i + Z_i \tag{6.30}$$

then the still water equilibrium is a stationary solution of the scheme (6.17),(6.20).

Proof of Theorem 6.1. We have defined the corrected fluxes (6.20) to preserve the flat water level ($h_i^n + Z_i = cst$) and using (6.26),(6.29) we can write

$$q_i^{n+1} = q_i^n - g \sum_{j \in K_i} \sigma_{ij}\, \tilde{h}_{ij} (\bar{h}_{ij} + Z_{ij} - (h_i + Z_i))\, n_{ij}. \tag{6.31}$$

So with the property (6.30), if $u_i^n = 0$, we have $u_i^{n+1} = 0$. □

Moreover we prescribe $Z_i = Z_j \Longrightarrow Z_{ij} = Z_i$ to recover the standard scheme for a flat bottom. Then a simple choice for Z_{ij} is

$$Z_{ij} = \sqrt{\frac{(h_i + Z_i)^2 + (h_j + Z_j)^2}{2}} - \bar{h}_{ij} \tag{6.32}$$

7 2D Numerical Results

The 2D scheme has been first tested on the same problems as the 1D scheme in a channel with constant width. The results are completely analogous to those in 1D so they are not presented here.

The second test case concerns a partial failure of a dam in a 200 x 200 m basin as depicted in Fig: 7.1. This problem has been computed by many authors (see e.g. [1]). The initial water height is $h_1 = 10.m$ on the left hand side and $h_1 = 5.m$ on the right of an idealized dam, the initial velocity is 0. Water is released into the downstream side through a breach 75. m wide. The discretization step of the boundaries is 2.5 m, the mesh contains about 8100 nodes and 15800 triangles. In Fig: 7.2 we show the watersurface elevation at time $t = 7.2s$. The velocity field is plotted in Fig: 7.3. On this fine mesh, the vertices on each side of the breach are accurately computed even with this first order scheme.

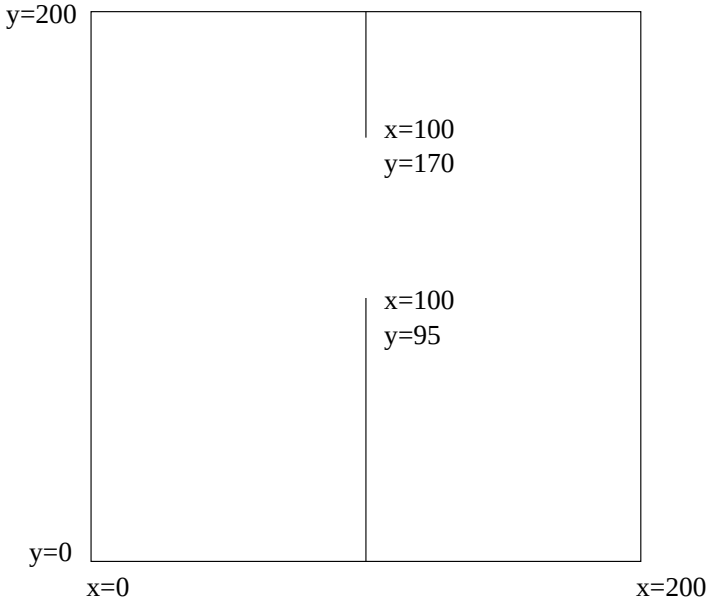

Figure 7.1: Geometry of the basin with dam breach.

We have also considered problems on a geometry given as test case in the code *Telemac* developed at EDF ([15]). It is a river with obstacles, the length is about 1,200 meters, the width 300 m, the bottom is varying from 252 to 260 m as shown

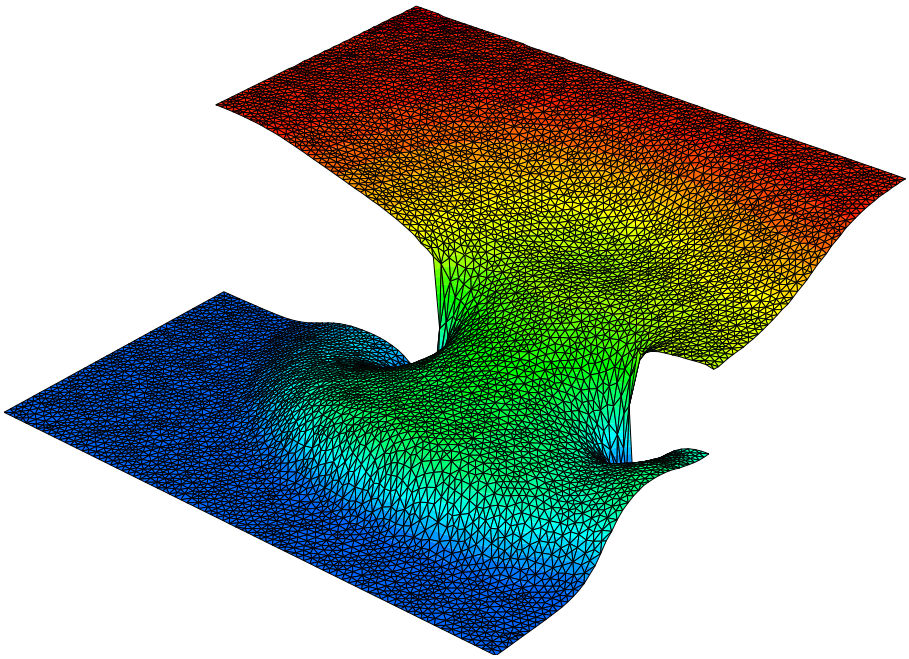

Figure 7.2: Partial dam break: Water elevation at $t = 7.2s$.

in Figure (7.4). When the water level is more than 260 m we assume that there are walls on the river sides. The mesh is about 1,000 nodes and 1,850 triangles. If we prescribe as initial conditions a constant water level and a velocity equal to 0. and an inflow discharge equal to 0, integrating in time we verify also on this geometry that the initial conditions are preserved up to the machine accuracy.

For another test, the initial water level is 265 m, and the inflow discharge is growing from 0 to $300 m^3/s$ for t varying from 0 to 1,000s. The CFL number used to define the time step is 1. The Figure (7.5) (resp.(7.6)) presents the water height (resp. the velocity field) at the final time.

Finally we have tested on this geometry the arrival of the water on a dry area. For this computation, we add a friction term to the water height equation which becomes

$$\frac{\partial h}{\partial t} + \mathrm{div}(hu) = -g\frac{q|q|}{K^2 h^{7/3}} \tag{7.1}$$

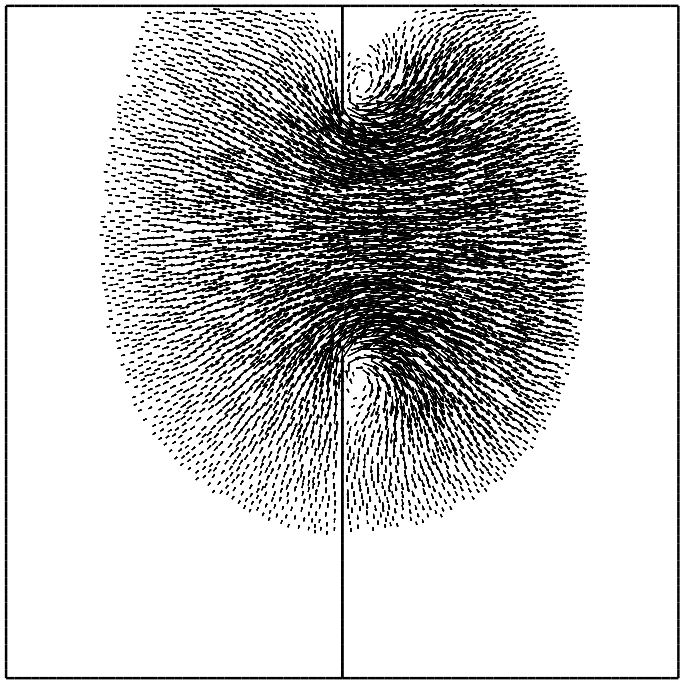

Figure 7.3: Partial dam break: Velocity field at $t = 7.2s$.

with K the Strickler coefficient (we use here $K = 25$). This source term is treated implicitly (see, e.g., ([21]). The initial condition is the river with a dry bed. Then we maintain as inflow boundary condition a water level equal to 254 m and the Figures (7.7)-(7.9) show the filling up of the river at times $t = 250s, t = 500s, t = 1000s$. On this example on dry varying bottom, we verify that the scheme preserves the water height positivity, so no clipping is necessary and the mass conservation is satisfied.

8 Conclusion

Following the theory of kinetic schemes for gas dynamics ([22]),([23]), we have applied these schemes to the Saint-Venant equations. There are two main difficulties. Firstly, the water height can vanish (the system loses strict hyperbolicity) and this is naturally handled by kinetic schemes. Secondly, finite volume schemes do not preserve the still water equilibrium. In this paper, our main target has been the

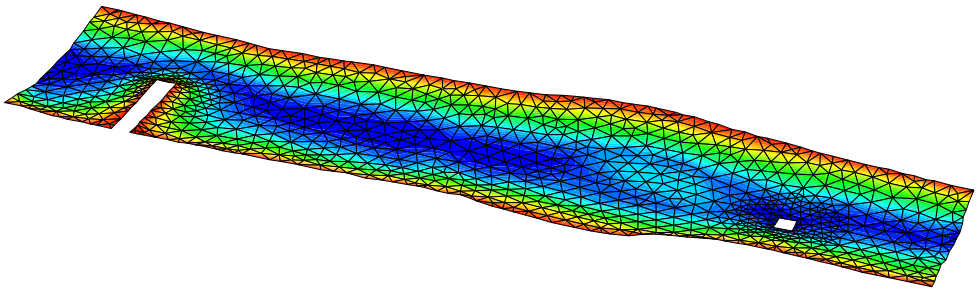

Figure 7.4: Bottom level.

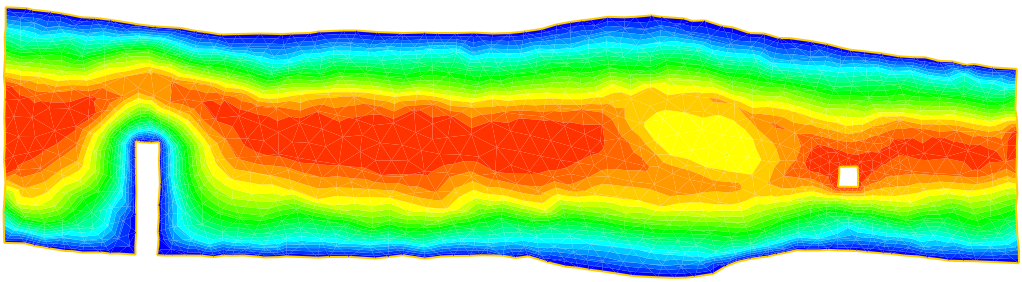

Figure 7.5: Water height.

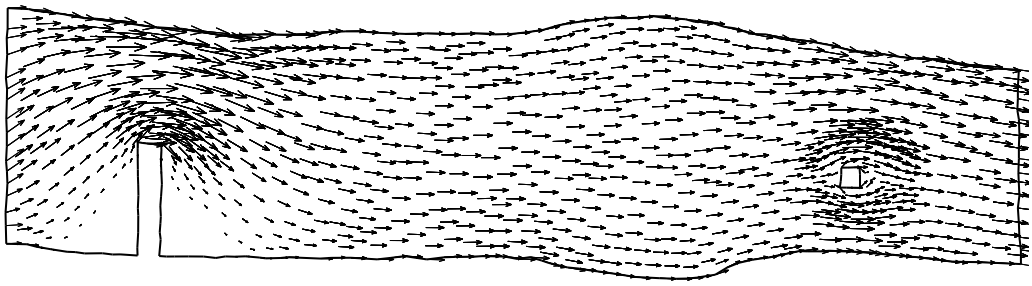

Figure 7.6: Velocity field.

definition of a modified scheme to deal with the source terms due to the bottom

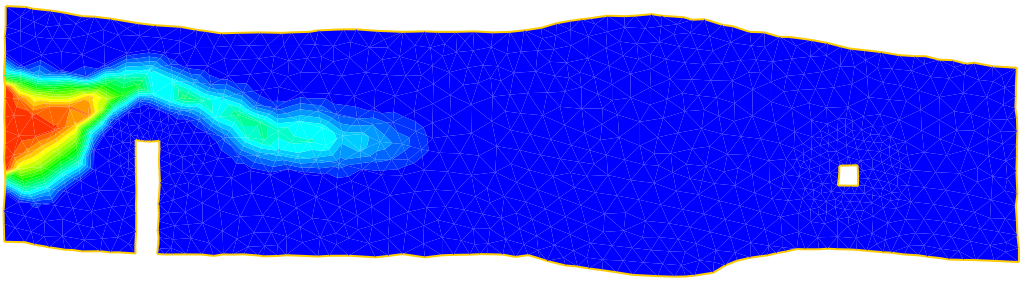

Figure 7.7: Water height at $t = 250s$.

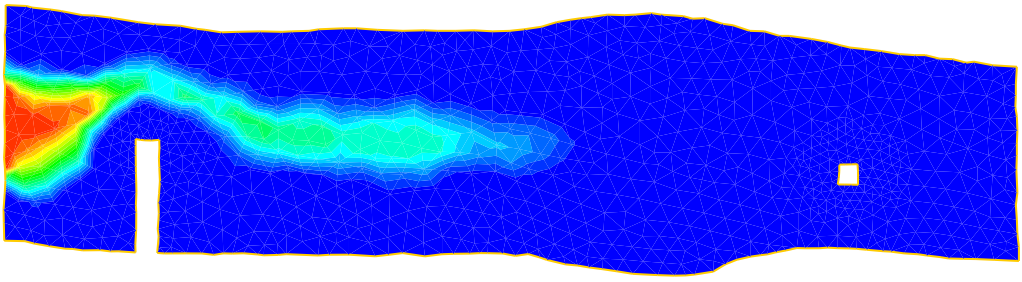

Figure 7.8: Water height at $t = 500s$.

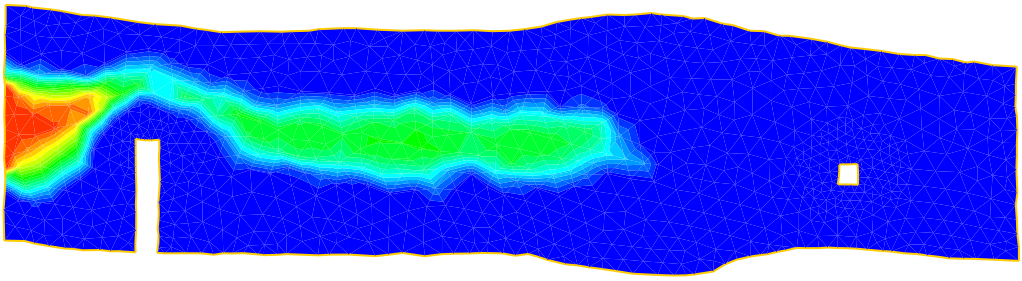

Figure 7.9: Water height at $t = 1000s$.

topography while preserving the still water equilibrium and the stability properties of the initial scheme.

We have not discussed here the treatment of the boundary conditions, it will be done in a forthcoming paper. Besides, to improve the accuracy, a second order scheme is under development, the difficulty is then to write a scheme preserving also the properties of the first order one which is not the case for a simple MUSCL like extension.

Acknowledgements The authors would like to thank M. Guesmia and J.M. Hervouet for helpful comments and suggestions.

References

[1] Alcrudo F. and Garcia-Navarro P., A High-Resolution Godunov-type Scheme in Finite Volumes for the 2D Shallow-water Equations, *Int. J. for Numerical Methods in Fluids*, Vol 16, 489–505 (1993)

[2] Angrand F., Dervieux A., Boulard V., Periaux J. and Vijayasundaram G., Transonic Euler simulation by means of finite element explicit schemes,*AIAA-83*, (1984).

[3] Bermudez A., Dervieux A., Desideri J.A., Vazquez M.E., Upwind schemes for the two-dimensional shallow water equations with variable depth using unstructured meshes, *Comput. Methods Appl. Mech. Engrg.* 155 (1998), no. 1-2, 49–72.

[4] Bermudez A., Vazquez M.E., Upwind methods for hyperbolic conservation laws with source terms, *Comput. Fluids* 23 (1994), no. 8, 1049–1071.

[5] Botchorishvili R., Perthame B. and Vasseur A., Equilibrium Schemes for Scalar Conservation Laws with Stiff Sources, *Inria report*, RR-8931 (2000)

[6] Bouchut F., Construction of BGK models with a family of kinetic entropy equations for a given system of conservation laws. *J. Stat. Phys.*, 95(1/2) (1999), 113–170.

[7] Bristeau M.O. and Perthame B., Transport of pollutant with kinetic schemes. In preparation.

[8] Dafermos C.M., *Hyperbolic conservation laws in continuum physics*, Springer Verlag, GM 325 (1999).

[9] Godinaud G., LeRoux A.Y., LeRoux M.N., Generation of new solvers involving the source term for a class of nonhomogeneous hyperbolic systems. Preprint.

[10] Godlewski E. and Raviart P.-A., *Numerical approximations of hyperbolic systems of conservation laws*, Applied Mathematical Sciences 118, Springer-Verlag, New York (1996).

[11] Gosse L., A priori error estimate for a well-balanced scheme designed for inhomogeneous scalar conservation laws, C.R. Acad. Sc. Paris 327 (1998) 467 – 472.

[12] Gosse L., Leroux A.-Y., A well-balanced scheme designed for inhomogeneous scalar conservation laws, C. R. Acad. Sc., Paris, Sér. I **323** (1996) 543–546.

[13] Goutal N. and Maurel F. Proceedings of the 2nd workshop on dam-break simulation, Note technique EDF, HE-43/97/016/B (1997).

[14] Greenberg J.M. and Leroux A.-Y., A well balanced scheme for the numerical processing of source terms in hyperbolic equations,*SIAM J. Num. Anal.* 33, 1–16 (1996).

[15] Hervouet J.M., Présentation du système TELEMAC, Note technique EDF, HE-43/96/039A (1996).

[16] LeVeque R.J., *Numerical Methods for Conservation Laws, Lectures in Mathematics*, ETH Zurich, Birkhauser (1992).

[17] LeVeque R.J., Balancing source terms and flux gradients in high-resolution Godunov methods: the quasi-steady wave-propagation algorithm. *J. Comput. Phys.* 146 (1998), no. 1, 346–365.

[18] Lions P.L., Perthame B. and Souganidis P.E., Existence of entropy solutions for the hyperbolic systems of isentropic gas dynamics in Eulerian and Lagrangian coordinates, *Comm. Pure Appl. Math.*, **49** (1996), no. 6, 599–638.

[19] Lions P.L., Perthame B. and Tadmor E., Existence of entropy solutions to isentropic gas dynamics system in Eulerian and Lagrangian variables, *Comm. Math. Phys.* 163 (1994) 415–431.

[20] Monthe L.A., Benkhaldoun F., Elmahi I., Positivity preserving finite volume Roe schemes for transport-diffusion equations,*Comput. Methods Appl. Mech. Engrg.* 178 (1999), 215-232.

[21] Paquier A., Modélisation et simulation de la propagation de l'onde de rupture de barrage, Thèse Univ. J. Monnet, 1995.

[22] Perthame B., An introduction to kinetic schemes for gas dynamics, *An introduction to recent developments in theory and numerics for conservation laws.* L.N. in Computational Sc. and Eng., 5, D. Kroner, M. Ohlberger and C. Rohde eds, Springer (1998).

[23] Perthame B. and Qiu Y., A variant of Van Leer's method for multidimensional systems of conservation laws,*J. Comp. Phys.* 112(2), 370–381 (1994).

[24] Perthame B. and Simeoni C., in preparation.

[25] Roe P. L., Approximate Riemann solvers, parameter vectors and difference schemes,*J. Comp. Phys.* 43, 357–372 (1981).

[26] Russo G., personal communication.

[27] Serre D. , *Systèmes hyperboliques de lois de conservation, Parties I et II.* Diderot, Paris (1996).

[28] Stoker J.J., The formation of Breakers and Bores, *Com. on Appl. Math.*, Vol 1, No 1 (1948).

[29] Van Leer B., Towards the Ultimate Conservative Difference Schemes. V. A Second Order Sequel to the Godunov's Method, *J. Comp. Phys.*, 32 (1979).

Unit é de recherche INRIA Lorraine, Technop ôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS L ÈS NANCY
Unit é de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unit é de recherche INRIA Rh ône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unit é de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unit é de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
http://www.inria.fr
ISSN 0249-6399